

RUNNING HEAD: DIGITAL SELECTION PROCEDURES

Personnel Selection in the Digital Age: A Review of Validity and Applicant Reactions, and
Future Research Challenges

Stephen A. Woods
University of Liverpool Management School

Sara Ahmed
University of Surrey

Ioannis Nikolaou
Athens University of Economics & Business

Ana Cristina Costa
University of Bradford
&
Neil Anderson
University of Bradford

ARTICLE ACCEPTED FOR PUBLICATION IN *EUROPEAN JOURNAL OF WORK AND
ORGANIZATIONAL PSYCHOLOGY* (DOI: 10.1080/1359432X.2019.1681401)

Address for Correspondence:
Stephen A. Woods PhD
University of Liverpool Management School
University of Liverpool
Chatham Street
Liverpool
L69 7BZ
E: Stephen.a.woods@liverpool.ac.uk

Abstract

We present a targeted review of recent developments and advances in digital selection procedures (DSPs) with particular attention to advances in internet-based techniques. By reviewing the emergence of DSPs in selection research and practice, we highlight five main categories of methods (online applications, online psychometric testing, digital interviews, gamified assessment and social media). We discuss the evidence base for each of these DSP groups, focusing on construct and criterion validity, and applicant reactions to their use in organizations. Based on the findings of our review, we present a critique of the evidence base for DSPs in industrial, work and organizational psychology and set out an agenda for advancing research. We identify pressing gaps in our understanding of DSPs, and ten key questions to be answered. Given that DSPs are likely to depart further from traditional non-digital selection procedures in the future, a theme in this agenda is the need to establish a distinct and specific literature on DSPs, and to do so at a pace that reflects the speed of the underlying technological advancement. In concluding, we, therefore, issue a call to action for selection researchers in work and organizational psychology to commence a new and rigorous multidisciplinary programme of scientific study of DSPs.

Keywords: Selection; Digital Selection Procedures; Online assessment; Validity; Applicant reactions; Internet-based selection

Personnel Selection in the Digital Age: A Review of Validity and Applicant Reactions, and Future Research Challenges

The rapid advancement of the digital economy has led to the emergence of an extensive array of new and innovative technological methodologies to assist with personnel selection decisions (McCarthy, Bauer, Truxillo, Anderson, Costa, & Ahmed, 2017; Nikolaou, 2014). Such is the scale of this advancement that these methods increasingly diverge from conventional methods of selection, requiring industrial, work and organizational (IWO) psychologists to establish an evidence base for their application. However, unlike advances in previous selection methodologies, digital selection procedures are subject to rapid innovation and development, potentially creating a widening gap between practice and the evidence base. To assist in gauging the extent of this gap, we present a targeted but comprehensive review of what we currently know from the literature on digital selection procedures, critically examining existing evidence gaps, and setting out a number of implications and directions for how these gaps can be closed. Although more general reviews of applicant perspectives and reactions in selection have been published (e.g. McCarthy et al., 2017; Nikolaou, Bauer, & Truxillo, 2015; Truxillo, Bauer, McCarthy, Anderson, & Ahmed, 2018), our review concentrates upon research and practice in digital selection procedures (DSPs) for organizational staffing.

In this review, we examine the literature on DSPs in three main sections. In the first, we review emergent trends in digital selection procedures that have gained traction in organizational human resource management practices. We set out the evidence base from applied psychology concerning issues of validity in digital assessment methods around the measurement and predictive robustness of DSPs, and applicant reactions to DSPs. In the second section, we present a focused, critical appraisal of the evidence base and highlight the

main substantive gaps in our current understanding of DSPs in research. In the third section, we outline an agenda for research to address these gaps, highlighting six core aspects that we propose should frame future research on DSPs. Critically, these reflect research issues that are especially salient in the digital age, rather than simply application of past research challenges in selection.

In our review, we focus on the main research published post-2010, given the extent of change and development in digital technology during the recent past. However, where relevant we highlight key studies published before this decade. We also focus on research from IWO psychology, with an aim to provide commentary and appraisal of research in the discipline specifically. While research on digital innovation may appear from a variety of sources, our interest was to understand specifically the nature of research in the IWO psychology scientific literature.

We make three substantive contributions to the literature on DSPs. First, we provide a timely, updated, and targeted review of what we know about DSPs currently, drawing together a hitherto disparate evidence base. Second, we use this review to set out a call for a new and rigorous research effort on the part of IWO psychologists to develop a distinct literature on DSPs, addressing the unique challenges that they present in practice. Third, and finally, we set out prospective directions in that effort in six areas and present some example key questions that remain to be answered in each concerning DSPs.

Digital Selection Procedures: Main Applications and Emergent Evidence

The impact of technology on employee selection has been significant with the appearance of new technology-based methods or the digital adaptation of existing methods. Development of the technology and application of DSPs has been particularly rapid over the past decade, with organizations, in some cases, forging ahead to use various aspects of DSPs almost ahead of the evidence-base, or at least with little knowledge of very recent research

and recommendations for practice. That is, the assumption of ‘equivalence’ between traditional methods of candidate assessment and digitally-mediated methods has been quite prevalent, yet the evidence to support such assumptions has been somewhat less uniform or convincing. In this first section, we bring together this evidence, but also draw links between research and practice in this rapidly changing, technologically-driven area of IWO psychology practice in organizations internationally.

We define DSPs as any procedure that makes use of digital communication technology (i.e. computer-, internet- or mobile-based) for the purposes of assisting organizations during recruitment and selection. Trends in the literature on development and applications of DSPs have converged around five main types of applications in organizations and human resources. These are online applications, psychometric testing, digital interviews, gamified assessments, and the use of social media as a source of assessment data. In IWO psychology, selection methods tend to be evaluated from two perspectives, namely validity and applicant reactions (Moscoso, Salgado, & Anderson, 2017). We review literature in this section through these two lenses, examining evidence around these five main emergent clusters of DSPs.

Online Applications

Online applications are standardized online forms where applicants are required to provide personal details as well as information relevant to the job for which they are applying. The use of online applications for the purpose of screening applicants increases efficiency by shortening hiring cycles, leading to a significant reduction in selection costs, while simultaneously expanding applicant pools (Bauer et al., 2006). Online-applications have also shown to be more effective in assuring objectivity in the handling of job applications and reducing the potential for adverse impact for protected groups (Konradt, Warszta, & Ellwart, 2013). However, peer-reviewed literature reveals little on the validity of online applications. Like non-digital applications, data collected is more often declarative

(i.e. educational and occupational history) rather than construct-driven. Such data are therefore unsuitable for construct-based scoring or validation.

Online applications are often the first impression job seekers gain about an organization and its HR policies, underlining the importance of applicant reactions to employee selection processes. Applicant perceptions are related to applicants' organizational attraction, behaviour, decisions to drop out, and attitudes toward employers and organizations (e.g., Konradt et al., 2013; Truxillo et al., 2018). Early studies provided encouragement about the positive perceptions of applicants about online-applications (see Bauer et al., 2006; Lievens & Harris, 2003). However, features of the system are important with justice perceptions shown to be related to the design, content and user-friendliness of the interface (De Goede, Van Vianen, & Klehe, 2011; Selden & Orenstein, 2011) and degree of customization to the applicant (Kraichy & Chapman, 2014). Studies conducted prior to the last decade showed that characteristics of the applicant determine or moderate perceptions of online applications. These included the applicant's age, his/her internet knowledge and exposure (Sinar, Reynolds, & Paquet, 2003), and person-organization fit (Braddy, Meade, Michael, & Fleenor, 2009; Dineen, Ling, Ash, & Del Vecchio, 2007; Pfieffelman, Wagner, & Libkuman, 2010). As interaction with the digital world is now the norm for most industrialized societies, factors relating to familiarity or exposure to the internet may be increasingly irrelevant among the working-age population.

Moreover, some organizations have incorporated Artificial Intelligence (AI) in their online application system in which the system selects the most qualified applicants. This can provide timely automated feedback to the applicants and save time and cost for the organization. Recent research conducted among applicants for public school teaching positions (Sajjadi, Sojourner, Kammeyer-Mueller, & Mykerez, 2019) illustrated initial evidence that AI can be used to scan online job application and evaluate the relevance

candidates' work experience, history of tenure and voluntary and involuntary turnover, and history of approaching jobs, and translate them into predictor of performance and work outcomes. Van Esch, Black and Ferolie (2019) found that applicants' attitudes toward the use of AI in their online application process had a positive influence on their probability of completing the application.

Psychometric Testing

In our review, we focus on the digital presentation of tests through the internet (online testing), and more recently through mobile devices. This differs to simple computerized tests of course in that applicants typically log-on remotely via an internet connection and may complete non-proctored version of tests (Ryan & Derous, 2019; Scott, Bartram & Reynolds, 2018).

Two pressing areas identified early in the evolution of internet testing have dominated the research literature concerning measurement equivalence (Potosky & Bobko, 2004; Ployhart, Weekley, Holtz, & Kemp, 2003) and test security (including test-taker integrity versus cheating; Naglieri et al., 2004), issues which also apply to mobile-delivered tests (Morelli, Mahan & Illingworth, 2014). Measurement equivalence for online tests may be considered across proctored and unproctored conditions as well as internet and paper-and-pencil formats (Beaty, Nye, Borneman, Kantrowitz, Drasgow & Grauer, 2011; Tippins et al., 2006). Studies examining the equivalence of paper-and-pencil versus online tests have generally focused on noncognitive tests such as personality inventories, yielding mixed findings ranging from general measurement invariance on the one hand (e.g. Meade, Michels and Lautenschlager, 2007), to indications of test-format effects on the other (e.g. Coyne, Warszta, Beadle & Sheehan, 2005).

However, methodological limitations of equivalence studies may bias results. Le Corff, Gingras and Busque-Carrier (2017) reviewed past studies and argued that separation of

intergroup and format effects was difficult because of the between-subjects design employed in most studies. In their study, they examine within-subjects completion of internet and paper-and-pencil formats of the same personality scales. Their data indicated close concordance of results based on quantitative and qualitative comparisons.

As digital engagement by test-takers has moved increasingly to mobile devices, so research has sought to examine construct validity and measurement equivalence in these formats. Arthur, Doverspike, Muñoz, Taylor and Carr (2014) showed measurement equivalence between Internet-based and mobile testing on a high-stakes test battery. They reported equivalence for both personality and cognitive tests, save for the observation of lower scores on cognitive tests completed on mobile devices. Brown and Grossenbacher (2017) also found measurement equivalence from completion on mobile and non-mobile devices, and also did not replicate the systematic score differentiation reported by Arthur et al. (2014). King, Ryan, Kantrowitz, Grelle and Dainis (2015) in further contrast, did not show measurement equivalence for a cognitive test completed on a PC versus a mobile device. In a more in-depth exploration, they proposed that mobile anxiety and potential higher distraction on mobile devices may be an issue. They also show that applicants felt less opportunity to perform on the mobile version.

Online tests also introduce variation into the standardized conditions under which tests are completed so issues of equivalence between forms become more complicated. Tippins et al. (2006) argue that this has potential effects on construct validity, especially in the case of speeded tests (e.g. of cognitive ability). However, Templer and Lange (2008) found measurement equivalence in both between-subjects and within-subjects comparisons of proctored and unproctored internet-based tests, concluding no effects of administration method on measurement validity.

The focus on measurement questions neglects, however, a critical concern in online testing from impersonation and fraudulent completion of tests in non-proctored settings (Lievens & Burke, 2011; Tippins et al., 2006). Research has considered ways of combatting such fraud (through for example analytics; Guo & Drasgow, 2010; Tendeiro, Meijer, Schakel & Meji, 2013; or remote proctoring; Karim, Kaminsky & Behrend, 2014). Alternative treatments of the issue in the literature have considered the role of process. For example, Lievens and Burke (2011) analysed data from field samples where a proctored verification test was used after the unproctored test and reported negligible evidence of cheating (detected through aberrant score differences). Moreover, Landers and Sackett (2012) argue that the advantages of attaining a wider applicant pool for a selection process through the use of unproctored testing could result in higher average performance gains (through e.g. setting higher cutoff scores), even if cheating occurs by some applicants.

Research pre-2010 indicated that applicants generally favoured internet-based testing over pen-and-paper (e.g., Ployhart et al., 2003; Potosky & Bobko, 2004; Salgado & Moscoso, 2003; Templer & Lange, 2008). No longer restricted to perform the online tests on their laptop or PC, applicants have also started to use their smartphones or tablets. Mobile testing seems particularly popular for applicants from particular demographic segments, such as younger applicants, women, Hispanics and African-Americans (Arthur et al., 2014; Morelli et al., 2014). Yet, comfort with testing mode may significantly affect applicant reactions and test performance. Indeed, a recent study by King et al. (2015) revealed that anxiety related to using mobile testing can affect applicants' test performance and lead to more negative reactions toward mobile testing. Also, environmental and context factor related to the use of mobile testing (e.g., motion, auditory, and visual distraction) can lead to lower test performance as well as lower efficiency and effectiveness when compared to PC internet testing (Coursaris, Hassanein, Head, & Bontis, 2012) leading to less favourable reactions

among job applicants. Questions remain about the reasons for these preferences. For example, Bruk-Lee et al. (2016) revealed favorable applicant reactions to the digitization of psychometric tests (e.g., 2D and 3D animation, and live-action video) as the digital format provided them with a greater opportunity to perform compared to the text-based format, engendering higher perceptions of fairness and job-relevance of the procedures as well as the organizations that use them.

Situational judgment tests. Situational judgment tests (SJTs) are traditionally text-based and are often completed online, but have also advanced quickly to make use of multimedia digital elements (Weekley, Hawkes, Guenole & Ployhart, 2015). Jackson, LoPilato, Hughes, Guenole and Shalfrooshan (2016) examined variance in scores on an online SJT, reporting that the main source variance in performance is the candidate ‘overall judgment effect’ (i.e. a single factor explaining judgment effectiveness). This replicates issues observed in using SJTs to measure multiple constructs (see e.g. Weekley et al., 2015, and Christian, Edwards & Bradley, 2010 for reviews). Arthur et al., (2014) reported properties of an online written SJT and observe that response formats affect construct validities, with more cognitively loaded formats (such as most/least response formats) being more strongly correlated with cognitive ability than personality, and Likert-rated responses being more strongly correlated with personality.

The presentation of detailed scenarios through video in SJTs has some potential advantages, such as more positive applicant reactions (Patterson, Ashworth, Zibarras, Coan, Kerrin & O’Neill, 2012). Moving SJTs to digital multimedia formats may also have an impact on the constructs being measured. For example, comparing written and video-based SJTs (identical tests, presented either on video or in writing), Lievens and Sackett (2006) reported that the video-based version had lower correlations with cognitive ability and was more predictive of interpersonal criteria.

The choice of the response mode (video vs text) can impact the validity and cognitive saturation of SJTs. In a field study conducted among police trainees, Lievens, De Corte, & Westerveld (2015) compared two different response formats of the same multimedia SJTs: behavioural video-based response (by recording applicants' response) and written response (where the candidates wrote their answer and submitted it online). They found that only the video-recorded response mode was a significant predictor of job performance and applicants perceived it more favourably in terms of communicating their reply than the written response mode.

Digital Interviews

The third group of DSP techniques are digitally-mediated interviews. At a basic level, interviews might be mediated through videoconference (see Levashina, Hartwell, Morgeson & Campion, 2014 for a review). However, more recent digital developments in interviewing depart from the interpersonal exchange that has typically characterized earlier techniques. These involve 'on-demand' interviews often during the initial steps of the selection process in which interviewees record their responses to a standard set of questions, which are then evaluated by either raters or computers through AI (Guchait, Ruetzler, Taylor & Toldi, 2014). Langer, König, and Krause (2017) compared interview ratings of video conference interviews with digital interviews and reported that ratings were generally higher for the latter. This may reflect the additional preparation time that interviewees have available before recording their answers. Ratings of digital interviews were also evaluated in a lab-based study by Gorman, Robinson, and Gamble (2018), who reported that ratings showed good internal consistency.

However, AI-based interviews have not been examined in published studies of construct validity, which may reflect their relative novelty and the proprietary nature of the technology that drives them. Yet, there is a widespread practice of the application of such technologies by HR technology providers like HireVue (see hirevue.com) and Montage (see

montagetalent.com). However, some of the technology on which they draw has been evaluated in the literature. One example is the analysis of digital interview responses through evaluation of transcribed text of answers to determine personality characteristics, drawing on research on language use and traits (see e.g. Schwartz et al., 2013; Tausczik & Pennebaker, 2010). People's use of language has shown to differentiate some personality characteristics, which could be useful constructs for selection assessment. AI-based evaluation of interviews may also attempt to measure wider indicators of the attributes of the applicant during the digital interviews, such as facial expressivity, seconds between responses, body temperature changes, word speed, and so on. Such indicators might, for example, be theorized to indicate emotional style or sensitivity. The purpose of capturing these and other indicators might be to develop algorithms that predict job performance potential from a multitude of different indicators. This next stage of digitization of selection interviews presents new challenges in evaluating their effectiveness, which we later discuss.

Studies of applicant reactions to digital interviews provide some cause for concern. Several studies reveal that applicants perceived video-conference interviews as less fair and favourable than face-to-face interviews, offering fewer opportunities to perform, having more difficulty regulating and understanding conversations compared to face-to-face interviews (Chapman, Uggerslev, & Webster, 2003; Sears, Zhang, Wiesner, Hackett, & Yuan, 2013; Silvester & Anderson, 2003; Straus, Miles, & Levesque, 2001). These findings indicate that when the selection procedure requires human interaction, such as in the case of interviews, using the Internet as a medium might hinder the transmission of body language or some verbal and nonverbal cues.

Studies on digital interviews showed that applicants perceived digital interviews as "creepier and less personal", induced more privacy concerns, led to ambiguity, lower fair perceptions, and less perceived controllability and social presence compared to video-

conference interviews (Langer et al., 2017, p371) and applicants' personality moderated their reactions (Brenner, Ortner, & Fay, 2016; Hiemstra, Oostrom, Deros, Serlie, & Born, 2019). In contrast to those findings, Suen, Chen, and Lu (2019) found no differences in perceived fairness between digital interviews that use AI or human rater and synchronize video-interviews among job applicants, although they exhibited lower favorability to digital interviews. These results indicate that applicants perceived computer and AI decisions as similarly trustworthy to humans, thereby not influencing their fairness perceptions (see Ötting, & Maier, 2018). However, as earlier highlighted, such findings may quickly become out of date as people become more comfortable interacting using digital media.

Gamified Assessments

The fourth category of DSPs centers around *gamified assessments* or *serious games*. Gamification refers to the use of gaming elements in non-game contexts (Deterding, Dixon, Khaled & Nacke, 2014; Armstrong, Lander, & Collmus, 2016a). In the context of assessment, games may be used in two ways. The first is using games as stand-alone assessments, the second is to enhance pre-existing assessments such as SJTs or personality questionnaires with game elements (Nikolaou, Georgiou & Kotsarilidou, 2019; Armstrong, Ferrell, Collmus & Landers, 2016b; Chamorro-Premuzic et al., 2016).

The use of game elements in the selection process might reduce faking, since desirable behaviours may be less obvious while playing a game, and as a result, improve the quality of information about applicants and prediction of job performance (Armstrong et al., 2016a). Using mobile or computing devices, candidates are exposed to a gamified environment or virtual world. Virtual worlds may be similar to real work settings and avatars may represent employees, for the purpose of eliciting job-relevant behaviours in situations similar to those taking place in a working environment (Laumer, Eckhardt, & Weitzel, 2012). Alternatively, virtual worlds or gamified assessments do not necessarily have to present realistic work-

related scenario, in order to reduce further the possibility of faking and social desirability. Moreover, the use of game elements in the selection process might promote fun, transparency, challenge and interaction. Games enable players to interact and compete with each other (Tippins, 2015), in team-based games or act alone in individually-based assessments. Game elements can also be applied to psychometric tests, for example to situational judgment tests to effectively assess candidates' soft skills. Surprisingly, despite much buzz about the use of gamified assessments in practice, there remains scarce published literature on the construct validity of gamified assessments and applicant reactions to them in the peer-reviewed IWO psychology literature. Only one recent study by Georgiou, Gouras, and Nikolaou (2019) examined the construct validity of a new gamified SJTs built to assess applicant's soft skills and found preliminary support that the construct validity of the gamified SJTs. This finding indicates that the addition of game elements (e.g., narrative and visual/voice-overs, avatar, and feedback) to the assessment method and its conversion into online format can be a valid and fruitful avenue for future research.

Social Media

The fifth and final category of DSPs focuses on the use of social media and network websites (SNWs), which could potentially be used by organizations to draw inferences about candidates' Knowledge, Skills, Abilities and Other Characteristics (KSAOs). Employers and professional recruiters can take advantage of the many new paths *Big Data* offers, such as using applicant "digital footprints" (e.g., professional-oriented profiles: LinkedIn, nonprofessional-oriented profiles: Twitter and Facebook, Aguado, Rico, Rubio, & Fernández, 2016; Nikolaou, 2014) as a potential source of information.

The use of social media for the purposes of selection assessment is often examined through the lens of ethical and privacy concerns (Brown & Vaughn, 2011; Stoughton, Thompson, & Meade, 2015), an issue we later return to. Many acknowledge that the practice

of viewing social media to assist in hiring decisions is widespread, and it has been suggested that most people expect (and are comfortable with) employers checking their social media profile, especially in professionally-oriented SNWs, such as LinkedIn (e.g. Vicknair, Elkersh, Yancey & Budden, 2010; Chamorro-Premuzic, Winsborough, Sherman, & Hogan., 2016). However, other reviews have urged caution in the use of such profiles and have raised a number of concerns over veracity, validity, personal information usage, compliance procedures, and other important issues (McCarthy et al., 2017; Ployhart, Schmitt, & Tippins, 2017).

From a construct assessment perspective, selection practitioners in organizations might use the content of profiles on SNWs to subjectively infer, or analytically derive job-relevant characteristics. Construct validity evidence for the use of social media depends on the means by which the digital data are evaluated. On a basic level, social media might be evaluated by a human assessor. In the case of personality traits, a common target construct assessed using SNWs content (Kluemper & Rosen, 2009, Kluemper, Rosen, & Mossholder, 2012), Van Iddekinge, Lanivich, Roth and Junco (2016) reported weak correlations between ratings of the social media content and measures of, for example, conscientiousness measured by conventional means. They reported similarly weak correlations for cognitive ability. Critically they also reported zero correlation between recruiter ratings of social media profiles and supervisor ratings of job performance, leading the authors to caution over the widespread usage of social media platforms in employee recruitment.

Adopting a different approach, Aguado, Andres, Garcia-Izquierdo and Rodriguez (2019) coded features of LinkedIn profiles of a sample of ICT employees (for example, the number of previous employment roles, the length of the description of experience, number of groups followed, or charity causes listed), and conducted exploratory factor analyses on the resulting coded data. They identified four factors, breadth of professional experience, social

capital, interest in knowledge updating, and breadth of non-professional information. They reported that these factors were predictive of productivity and work outcomes. For example, social capital was correlated with productivity in junior ICT workers and breadth of non-professional information correlated with absenteeism for senior employees.

The use of SNWs for selection purposes can have negative impacts on the applicant reactions (Madera, 2012), yet, this may vary between different age groups. A recent study by Nikolaou (2014) reveals that job applicants still use job boards more extensively than SNWs, in particular, younger applicants tend to use non-professional SNWs (i.e. Facebook) more, whereas older and male applicants were more likely to use professional SNWs (i.e. LinkedIn). Yet, applicants may also be less likely to have *faux pas* postings, especially if they are aware that a high number of organizations are using the information posted on SNWs as part of their assessment and selection (Roulin, 2014). They also are more likely to use impression management tactics in their social media to enhance their chances in getting the job they desire, which might also threaten the validity of using SNWs as a selection tool (Roulin & Levashina, 2016). Evidence of impression management in SNWs presents a serious concern regarding the accuracy of any evaluations drawn by recruiters in respect of selection assessment and job suitability. Furthermore, adverse impact on the profile ratings can present a potential bias issue among gender, age and minority groups. A recent study of SNWs in selection by Van Iddenkinge et al. (2016) revealed that women received significantly higher ratings than men and that minority received lower ratings for overall suitability, but not for evaluation of KSAOs.

A Critical Appraisal of the Evidence Base for Digital Selection Procedures

There are evident advantages to employing digital technology in selection procedures. Such technologies make selection procedures faster, easier, sometimes more vivid and fun while expanding the number of applicants by reducing barriers of distance, cost, and time.

However, our review of the evidence base in industrial, work and organizational psychology highlights significant limitations in our scientific understanding of their effectiveness. In this section, we provide a critical overview of this evidence base, before proceeding to set out a research agenda for closing the gaps we identify.

Validity

The increasing use of digital methods of selection assessment raises questions about how the transfer of measurement from conventional to digital techniques might affect validity. The issues are made more complex as digital forms of assessment depart further from conventional techniques. For example, hosting psychometric testing online rather than on paper might simply represent a presentational adaptation of the material. At a more complex level, digital test formats might allow content adaptation of tests through selection of items from item-banks or application of item response theory to compose individual-specific combinations of items. More complex still are the validity issues that surround completely novel forms of measurement such as gamified assessment or the use of algorithms to read digital data.

Alongside issues of construct validity is arguably the most critical gap currently in the literature on the validity of DSPs; namely the absence of peer-reviewed published studies of criterion validity. In the papers we reviewed, only two reported criterion-related validity of digital forms of assessment in the context of selection (Van Idenkinngge et al., 2016; Nikolaou et al., 2019). The lack of criterion validity evidence is troubling for practice. Psychologists and other practitioners in Human Resources have for the past two decades had firm research foundations from which to make evidence-based selection methodology decisions. The meta-analysis of Schmidt and Hunter (1998) for example set out clearly the comparative criterion validities of different selection approaches. The digital revolution has introduced a raft of new techniques and methods, for which there exist little or no data on their predictive

properties with job performance and other criteria. The speed of the uptake of these methodologies in organizations risks rendering the conventional evidence base of criterion validity in selection increasingly less relevant.

Adverse Impact

Relatedly, research has largely neglected the issue of adverse impact in DSPs, with only the study of Van Idenkinngge et al. (2016) reporting analyses of subgroup differences in the ratings of candidates based on SNWs. Also, Lievens et al. (2015) found gender subgroup differences in using video SJTs when comparing the response mode, in which male candidates outperformed the females in the video-behavioural response mode and no such result was found in the written response mode. This study was conducted among police candidate, a male-dominated setting, which may explain the bias against female in the video SJTs setting. Moreover, a recent study (Suen et al., 2019) revealed that first impression and physical appearance influenced interviewers' rating in both digital (asynchronous) and video-conference (synchronous) interviews, however, digital interviews lessened this effect among human raters, indicating that digitization can help us reduce human bias.

However, the relevance of evidence in those studies is confined to evaluations that are made by human raters. The use of DSPs raises in addition, the significant issue of how bias may exist in algorithms that might be used for example to scrape web data or evaluate online applications and digital interviews. Studies from other fields have demonstrated the potential for bias in digital algorithms (e.g. Lambrecht & Tucker, 2019), illustrating the validity of wider concerns about their effect on society (Courtland, 2018; EU FRA, 2018). The perception that DSPs may somehow remove the bias from selection afforded by human raters (Kleinberg, Ludwig, Mullainathan & Sunstein, 2018) ignores the prospect that algorithms may simply replicate biases that exist already in society.

Privacy Concerns

A particularly challenging concern is information privacy in DSPs that can negatively impact applicant reactions to DSPs and consequently to the hiring organizations. Indeed, in the context of online screening, Bauer et al. (2006) found that applicants with higher privacy concerns had lower justice perceptions, which in turn, affected their intentions and attraction towards the organization. Particularly in the age of the DSPs, applicants may feel that organizations are invading their information privacy when using information from social networks, perhaps originally intended only for family and friends (Black, Stone, & Johnson, 2015). With the increased use of DSPs and the latest Facebook–Cambridge Analytica data abuse scandal, where data of about 87 million Facebook users were shared with Cambridge Analytica (see Griffen, 2018), many individuals are more concerned about the potential misuse of their personal and employment-related information submitted online during the selection process. This could of course also include illicit access of personal data held by organizations through data theft, digital security breach or hacking. As some organizations have started using candidate information on SNWs, applicants may perceive that as an invasion of their privacy. Indeed, Stoughton et al. (2015) in a laboratory setting used realistic selection scenarios to examine reactions to SNWs assessment. Participants were informed that the hiring organization had evaluated their social media to assess their professionalism. They found that SNWs screening had caused applicants to feel that their privacy had been invaded, resulting in increased intentions to litigate and lower organizational attractiveness. This is especially salient given that a recent national survey report within the U.S. among 2,380 hiring managers and HR professionals across industries and company sizes in the private sector (Career Builder, 2017) revealed that 70 percent of the U.S. organizations are using social media to screen their applicants, up from 60 percent in 2016 and 11 percent in 2006, 57 percent are less likely to interview candidates that are not present online, and 54 percent rejected some applicants based on information on their social media. Some

organizations may even ask the applicants to provide their Facebook login information during the interview to perform a social media background check (Barnett, 2012).

Other organizations are hiring specialized companies to screen their job applicants by gathering all publicly available online content to draw out the personality of those applicants using AI (e.g., FAMA, 2018a). They also claim that it is lawful and it can reduce the incidence of sexual harassment, prejudice, violence and other types of risk in the workplace without violating applicants' personal freedoms (FAMA, 2018b; Social Intelligence, 2018). However, using SNWs, especially non-professional SNWs, and AI to screen and evaluate applicants is associated with many concerns, such as raising potential bias and discrimination issues among minority groups, evoking applicants privacy concerns when using their SNWs' information without consent, which in turn, can lead to legal and ethical challenges (Drouin, O'Connor, Schmidt, & Miller, 2015; Roth, Bobko, Van Iddekinge, & Thatcher, 2016; Van Iddekinge et al., 2016). Indeed, a recent study by Becton, Walker, Gilstrap, and Schwager (2019) showed that using applicants' information on their non- professional SNWs negatively influenced recruiter ratings of applicants, whereas no effects were found when using their information on professional SNWs. This questions the validity (job relevance), and reliability of such information, and further raise fairness and privacy concerns (Brown & Vaughn, 2011; Clark & Roberts, 2010).

In response to these, and other, issues, a number of countries have established new privacy laws. For example, the European Union introduced the General Data Protection Regulation (GDPR) in May 2018 in which there is one set of data protection rules for all organizations operating in the EU, offering individuals more protection and control over their personal data, reshaping the way corporations approach data privacy (EU GDPR, 2018). With regard to using applicants' information on their SNWs, there are different views. In some countries, it is illegal to use applicants' information in SNWs to select them and consider it as

a breach of their privacy (e.g., Germany: Leggatt, 2010; and some states in the US: Stinson, 2014), while in other countries, the law is not so severe, or even no legal law exists with regard to this issue in other countries (e.g., Saudi Arabia: Anderson, Ahmed, & Costa, 2012; Ahmed, Anderson & Costa, 2014). This issue has attracted media attention, entailing several issues that employers should consider when using DSPs in order to avoid ethical and legal implications.

Digital Familiarity

Another issue related to applicant reactions to DSPs is the extent to which applicants are familiar with the new procedures. This might be more prominent in novel DSPs, such as digital interviews and gamified assessment. For example, a study by King et al. (2015) illustrated that applicants had more positive perceptions (i.e., chance to perform and test ease) of personal computers (PCs) internet testing as they were more familiar with it compared to the mobile version of the same test. Also, applicant age, experience with the DSPs are more likely to influence their reactions and test performance. This can also be more salient in some occupations where DSPs might be less job-relevance. For example, face-to-face interviews are essential in medical or academic selection to assess the suitability of the candidates since communication, presentation, and interpersonal skills are important in these positions and cannot be easily replaced with digital interviews (Zibarras, Patterson, Holmes, Flaxman, & Kubacki, 2018), whereas engineering, technical, and gaming industries can easily replace the traditional testing with the digital and gamified format, as experience with technology is an essential part of the job and would probably gathered positive reactions and improve the organizations' image. Also, as the global digital divide and differences in exposing to the latest technology vary across different countries —especially in developing countries, applicant familiarity can be a major issue (Ahmed, Anderson, & Costa, 2018). It can

negatively influence applicant reactions to DSPs, discouraging high- quality applicants from applying, or even withdrawing their application in those counties.

A Research Agenda for Digital Selection Procedures

In summary, our review of the literature highlights significant shortcomings in the scientific evidence base in IWO psychology surrounding the effectiveness of DSPs. To address these gaps, applied psychologists could be called upon to embark on a rapid and extensive programme of research on DSPs. In the remainder of this paper, we outline a possible research agenda to help focus that effort in key areas. Importantly, these areas reflect issues that are pertinent, especially to selection using digital procedures, rather than simply applying past research concerns in selection using non-digital methods. Our proposed agenda, therefore, aims to prompt the development of a new and specific literature on DSPs. We structure our proposals into six core aspects; 1) understanding construct validity of DSPs; 2) creating a literature on validity that promotes improvement and development of DSPs; 3) evaluation of the criterion-related validities of DSPs; 4) understanding adverse impact and bias in DSPs; 5) development of current and up-to-date perspectives on applicant reactions; 6) understanding the impact of privacy concerns for applicants and their potential implications for organizations. We also present ten related questions for research in the field (Table 1).

Construct Validity of DSPs

Empirical studies of the construct validity of digital assessments have, in the past, considered the issue from the perspective of establishing measurement equivalence with non-digital methods (e.g., Anderson, 2003; Breugh & Stake, 2000; Ployhart et al., 2003; Potosky & Bobko, 2004; Salgado & Moscoso, 2003; Templer & Lange, 2008). To the extent that psychological tests, for example, demonstrate measurement equivalence, then there is some confidence to be gained in terms of generalising findings from non-digital methods (e.g.

Schmidt, Oh & Shaffer, 2016). However, this is not, in our view, an acceptable long-term substitute for primary studies and an evidence base of the validities of DSPs. Rather IWO psychologists will need to move beyond seeing DSPs as digital *equivalents* of non-digital methods, to rather see them as techniques in their own right.

Focusing on measurement equivalence implicitly this affords primacy to non-digital selection procedures, holding out digital methods in comparison to this benchmark. Yet this comparison may be a flawed index of the measurement value of digital assessment for two reasons, illustrated by considering the case of psychometric testing.

Firstly, it is assumed that the paper-and-pencil administered form of the assessment necessarily remains valid, reflecting in part its standardised administration. However, one key variant is the respondents themselves. It is plausible that for people more used to interacting with digital rather than paper-based media, a paper-and-pencil administered assessment is sufficiently unfamiliar to undermine the properties of non-digital formats. The digital presentation could, in effect, be psychometrically superior. Secondly, in seeking measurement equivalence, there is no examination of whether the digital presentation of an assessment adds incremental criterion validity. That is, perhaps differences reflect additional information captured rather than inaccuracy or error.

The solution to these applied problems of measurement equivalence studies is to approach the construct validation of digital assessments in their own right, rather than benchmarking against non-digital formats (for example see Aguado et al., 2019; Gorman et al., 2018; and Arthur et al., 2014). This will require a departure from our assumptions of how construct validation might be approached. For example, rather than test items, analyses may be conducted on digital data captured from a variety of means (e.g. responses in a game, use of language, micro-expressions in interview and so forth). Those digital data might even be more directly reflective of performance constructs (e.g. online evidence of productivity or

networking; performance in a simulated work environment). In sum, research must show what assessments *are* measuring, instead of examining what they *may not* be measuring.

Building on this idea, we propose that in understanding construct validity in digital assessments, research should seek to understand the place of DSPs in a Multi-Trait-Multi-Method (MTMM) approach to construct measurement in selection. In a conventional MTMM paradigm, job-related constructs are established the ‘what’ of measurement, with methods addressing the ‘how’ (Ployhart, 2006; Woods & West, 2019). Digital methods represent alternative ways of assessing the constructs of interest to selection practitioners. Their construct validity might be viewed therefore in the context of what they contribute rather than on their functional equivalence of other methods. Staying with the example of personality trait measurement, analyses of alternative methods of assessment such as self-other agreement in ratings collected using surveys, are analogous to digital/non-digital methods (e.g. self-report and gamified assessment). Indicators of the construct validity of the measures in the former case (one self-perceptual, the other observation-based) include convergence to a reasonable degree. The same kind of evidence base may be established for the latter. The example of personality measurement is especially relevant because it has been argued that observer-ratings of traits improve criterion validities with job performance (Oh, Wang & Mount, 2011). The same may be argued for DSPs as part of an overall measurement strategy in selection. A challenge for researchers and practitioners is to understand how they enhance rather than replace existing techniques.

Applying Validity Data to Improve and Develop DSPs

Development of digital services and technology follows a radically different pathway from the traditional development of assessment methodology in IWO psychology. Digital technology is flexible and easily updated and adapted and so information from users, clients and others can be used to continually and rapidly improve the way that, for example, software

or online systems function. By contrast, collection and publication of validation data in the IWO psychology literature is a slower process, leading to relatively long lead time for empirical evaluation to impact practise or lead to the improvement of procedures. To illustrate, consider the process of design and validation of psychometric tests used in selection.

The history of psychometric testing in selection is built on foundations of classical ‘instruments’; tests that were well established in practice and supported by robust research. The lead time for the development of psychometrics required extensive trialling and data collection, and the production of materials in hardcopy indicated the expected longevity of published instruments. This approach prompted much criticism of the field’s capacity to innovate (see Grucza & Goldberg, 2007).

Digital psychometric assessment by contrast enables a substantial reduction in development time. For example, platforms such as MTurk and Prolific can provide access to psychometric trial data in a matter of days. The online presentation of materials also means that there is no compelling reason for psychometrics to be conceptualised as fixed instruments. Rather, individual components of occupational tests can be compiled in modules and adapted to individual assessment contexts (Woods, 2018). In the area of personality assessment, for example, items and scales can be reconfigured to create parcels or composites directed to performance demands (such as competency scoring in personality inventories). Data on the comparative properties of personality scales may guide such reconfiguration (e.g. Woods & Anderson’s 2016 *Periodic Table of Personality*).

The rapid configurable nature of digital assessments means a fundamental shift in the way we approach validation, from an ‘endpoint’ of instrument creation to an ongoing accumulation of insight into a technique or methodology. This implication extends beyond psychometrics to all forms of DSPs. Accumulation of live data from these forms of

assessment could be quickly evaluated to determine construct validity implications. It is then incumbent for psychologists to treat validation data, where shown to be weaker than desirable, as critical for improvement and development rather than as a generalized refutation of the effectiveness of a specific technique. Rather, by working with computer scientists and others, validity data can be used to quickly enhance and improve assessments for example or understand how changes in configuration might affect the underlying measurement properties. For this to be effective, IWO psychology must enable specific dissemination and publication routes for such empirical data. These might include specific forms of an article that focus on fast communication of data analyses in a concise and open format. This may also encourage sharing of methodologically robust, negative or null findings, which may be highly informative as DSPs continue to evolve.

Criterion Validity of DSPs

Our review highlights the acute need in the literature for studies of the criterion validity of DSPs. This need will only be addressed through researchers approaching the task of establishing the criterion validity of new forms of assessment with the same motivation and energy of the past. Classical meta-analyses of criterion validities of selection methodology have arguably served to reduce the need for basic empirical research on performance prediction in selection. This kind of research must be renewed with vigour because the landscape of digital HR assessment now bears little resemblance to that under which much of the validity evidence of the past has been established. As many observe, the future is now (Chamorro-Premuzic et al., 2016).

We propose that research in this line should seek to address two research gaps concerning the criterion-related validity of DSPs. The first concerns the lack of data on the prediction of performance by DSPs as stand-alone techniques. As earlier highlighted, DSPs should take their place alongside non-digital selection assessments so that their validity can

be placed in context. However, we would also raise a note of caution about the direct comparison of validities with the findings for non-digital methods accumulated and published in meta-analyses (e.g. Oh et al, 2016). This is because a proportion of data in such meta-analyses were collected historically (i.e. well before the digital transformation of our economies that we are currently experiencing). Both the context of performance prediction, and the nature of the criteria that are measured in organizations (the latter of which are likewise developing and changing in the digital economy) are not directly comparable to the past. In short, DSPs are ‘of their time’, and so any comparison with non-digital methods of the past should reflect application within contemporary organizational settings.

Relatedly, the second gap concerns data on the incremental validities of DSPs, when included alongside other, more traditional forms of selection procedure. To illustrate, consider the ways in which an organization might utilise digital interviews. They might replace face-to-face interviews, or be used as a pre-screen before face-to-face interviews (i.e. reducing the interviewee pool), or included alongside a face-to-face interview to provide additional selection data. The utility of these different options depends upon the incremental validity attained by combining the techniques. If both are equally valid, and no incremental variance in performance explained by combining them, then the replacement option is justifiable. If incremental validity is obtained by combination, then inclusion of both with a view to combining the findings to make a hiring decision is sensible. If the digital interview had weaker criterion validity, but was still stronger than an application blank, then its inclusion as a pre-screen makes sense. Data to support this decision making is missing in the literature. This leaves organizations to rely on practical benefits as a means for decision making. For example, digital interviews might be more efficient, less costly, and more sustainable. All are important benefits, but should be factored into decisions alongside validity evidence.

Adverse Impact and Fairness of DSPs

In their review of the history of discrimination research in 100 years of the *Journal of Applied Psychology*, Collela, Hebl & King (2017) draw pessimistic conclusions about the progress of discrimination research in applied psychology, which they felt has trailed too far behind societal trends. Recent discrimination research in selection has continued to debate and explore issues of adverse impact and differential validity of cognitive ability testing (see e.g. Roth, Le, Oh, Van Iddekinge & Robbins, 2017; Roth et al., 2014; Berry, Cullen, & Meyer, 2014; Berry & Zhao, 2015). While it is not our intention to challenge the value of the focus on this key selection assessment methodology, we do note that to date, there exists no commensurately detailed published exploration of the fairness and adverse impact of DSPs.

We find this absence both surprising and troubling, because digital technology has the potential to create new forms of bias and discrimination that we may not have anticipated. A UK Guardian article in 2017 highlighted the issue of discrimination in facial recognition technology used in policing. Algorithms were found to be substantially less accurate for black people compared to white people, creating bias and discrimination when applied within communities. This may reflect the biases inherent in the development of software code or ‘training’ of AI with unrepresentative samples of the population (Fernández & Fernández, 2019; EU FRA, 2019). There are obvious implications for digital interviewing as a DSP, which may include facial expression analyses that may prove to be differentially accurate across racial/ethnic groups. Another study relevant to recruitment and selection examined a seemingly inert issue of the distribution of an online job advertisement in the STEM industry (Lambrecht & Tucker, 2019). The advert was intended to be gender-neutral, yet when field-tested fewer women saw the advert than men. The reason proposed by the researchers was that in optimizing cost-effectiveness, the algorithm showed the advert to more men, because in digital advertising it is typically more expensive to show advertisements to women

(Lambrecht & Tucker, 2019). This is a further instance of how the application of digital technology might perpetuate inequality in job opportunities.

These are just two examples of new forms of bias and discrimination that researchers in IWO psychology must better understand. We suggest that emergent evidence of digital bias should prompt researchers to expand the horizon of discrimination research quickly, so as to avoid similar conclusions to Collela et al., (2017), being levelled in the future at our effort as research psychologists now, in adapting to the digital age.

Applicant Reactions to DSPs

Research has generally fallen behind the rapid shift in organizational usage of DSPs for employee selection, and this is particularly the case with regards to applicant perspectives. Many organizations are adopting DSPs without considering applicant reactions to such new techniques and are not aware of the major consequence of such reactions on the organizations, as such negative applicant experience can reflect negatively on the organizations and cost them a tremendous loss and their image as a good employer (Steiner, 2017). There are many platforms where applicants provide their online reviews about the organizations and their selection practices, such as Glassdoor (Van Hoya, 2013). Applicants' posts can significantly enhance or damage the organizational image and reputation as an employer, thus, organizations should always take into consideration applicant reactions to their selection procedures, especially when they implement new DSPs.

Beyond the reactions of applicants to specific DSPs, researchers may need to develop better frames of reference to understand the antecedent characteristics of digital techniques that lead to the formation of reactions. These might include features of the different platforms through which DSPs are delivered, and different types of media (e.g., gamified-, mobile-, avatar-based tests) that are employed. Design features of gamified assessments (e.g. ease-of-use, mobile hosting or the nature of games themselves) could similarly affect reactions.

Mechanisms of scoring and rating may be influential, with the differences between the automatic vs human ratings in digital interviews and their positioning in different stages of the hiring process. It is not clear yet how organizations and applicants perceive the removal of human interaction, which is usually an essential part of many jobs.

The use of digital footprint data and the analyses of data through algorithms and AI highlights a further area for future research; that this the importance of maintaining justifiable decisions. In the case of drawing information from SNWs, information processing by human raters may be unsystematic, meaning that decisions are inconsistent and inadequately grounded in evidence. Similarly, if algorithm-based decisions are not accompanied by transparent parameters, hiring managers may be unable to provide applicants with reasons for selection decisions, likely leading to negative reactions and serious moral, ethical, and legal challenges (Zerilli, Knott, Maclaurin, & Gavaghan, 2018). In short, transparency of process and procedure in DSPs, if poorly understood, could foster negative perceptions of the practices (Gilliland, 1993). Research is needed to define for the digital age, the changing nature of *procedural* justice, and the associated antecedents and outcomes.

Privacy Issues in DSPs

As in the case of applicant reactions, research is needed to address fundamental questions about the the factors shaping and determining applicants' privacy concerns. This will help organizations when designing and implementing DSPs and associated privacy policies. For example, Black et al. (2015) introduced an updated model of the factors affecting applicant reactions to the use of SNWs as a selection tool. They posited that procedural factors, information factors, individual factors and socio-cultural factors will influence individuals' belief about the ability to control (or not control) SNWs information, which in turn will affect their perceptions of invasion of privacy and consequent behaviour (e.g., accept a job offer, recommend the organizations to others, and file a lawsuit). This

theory-driven framework illustrates how research can help advance our understanding of factors may influence applicants' perceptions of the invasion of privacy.

There are then further questions about the boundary conditions of such models. For example, generational differences in familiarity and preferences in DSPs can have a significant impact on applicants' privacy concerns. Reactions may vary across different occupations depending upon the relevance of such information to the job. For example, Jeske and Shults (2019) found that applicant perceptions of respect for privacy of social media screening varied across different jobs, in which respect of privacy was higher for jobs involved working with children (social services and childcare) or the government compared to sales and property management. There may be variation across countries and cultures, attributable to differences in cultural and contextual factors as well as privacy and data protection laws. The implications of research on these issues would help HR professionals to design their DSPs, improve the fairness and related privacy policies, ensuring effectiveness of DSPs across a wide range of contexts and situations.

Conclusions and Final Comments

Our review of the literature has highlighted significant developments in DSP research, implications for concomitant selection practices, and demonstrable need for future research into DSPs. However, it is also evident that the approaches and techniques that have been applied in conventional non-digital selection procedures need to be adapted to meet the demands of the digital format of assessments methods and the data that they produce. Our view is that the application of DSPs have forged ahead of scientific research and that in some areas organizations are using these new technologies rather 'blind' to their validity, adverse impact, privacy, or impact on applicants. Given the speed of advancement of some of the pertinent technologies this is perhaps unsurprising; what is now called for is a period of realignment between research and practice in IWO psychology in our view where these

concerns can be examined and addressed by applied research efforts into DSP in employee selection.

Many researchers have called for updated technology-based models of recruitment and selection (e.g., Black et al., 2015; Potosky, 2008; McCarthy, et al., 2017; Morelli, Potosky, Arthur, & Tippins, 2017) to conceptualize and create more specific, theory-driven hypotheses for technology now utilized in the personnel selection arena. Based on our review of the literature on DSPs we would go further, therefore, and extend a call to action to both researchers and practitioners to embark on a rapid and wide-ranging collaborative undertaking to establish a more rigorous and comprehensive evidence base for the application of DSPs.

All of the implications of our review point to the need for a new multi-disciplinary research programme on selection methodology for the digital age. In this programme, DSPs should take a specific place alongside conventional non-digital formats. For example, we should consider *digital interviews* to be a technique in their own right, rather than as an alternative for conventional interviews. Their application in practice should, like non-digital methods of the past, be based on evidence that they are psychometrically effective in the context of selection. As a result of such a programme, we may find that digital formats are superior replacements for methods of the past. Yet the utility of digital methods must be viewed beyond whether they outperform established techniques. There may be legitimate reasons, for example, why a digital assessment is considered more suitable by an organization than a non-digital one (e.g. it might create lower adverse impact for minorities, improved applicant reactions, or a better candidate experience), provided that criterion validity has been evidenced independently.

Addressing these needs presents a significant transitional opportunity for IWO psychology. The pace of development of digital technology is increasingly incompatible with

the slower processes of research dissemination in the field. The rise of DSPs is one of the most prominent early manifestations of this tension. Maintaining the centrality of our research evidence in the practice of selection will require changes to how we make available the findings of research to ensure they are relevant to state-of-the-art practice. But this issue is not peculiar to selection research, and rather is reflective of the impact of digital technology across business and management. Successfully meeting our *call to action* requires new learning by psychologists about the epistemologies, techniques and challenges of multi-disciplinary teams, and clarity about the benefits that psychologists can bring (i.e. the incentives for others to collaborate with us). However, by doing so and adapting to the digital age, selection researchers have an opportunity, as arguably in the past, to lead development of the tools, methods and processes of IWO research and practice, and ensure its continued impact in, and relevance to, organizations in the future.

References

- Aguado, D., Andrés, J. C., García-Izquierdo, A. L., & Rodríguez, J. (2019). LinkedIn" Big Four": Job Performance Validation in the ICT Sector. *Revista de Psicología del Trabajo y de Las Organizaciones = Journal of work and organizational psychology*, 35, 53-64.
- Aguado, D., Rico, R., Rubio, V. J., & Fernández, L. (2016). Applicant reactions to social network web use in personnel selection and assessment. *Revista de Psicología del Trabajo y de las Organizaciones*, 32, 183-190. doi.org/10.1016/j.rpto.2016.09.001
- Ahmed, S., Anderson, N., & Costa, A. C. (2014). Privacy and fairness perceptions of internet-based selection procedures. Poster presented at the 29th annual conference of the Society for Industrial and Organizational Psychology, Honolulu, Hawaii.
- Ahmed, S., Anderson, N., & Costa, A. C., (2018). Cross-Country Influence on Applicant Fairness Reactions to Internet-Based Selection Methods? Paper presented at the 29th annual conference of the International Congress of Applied Psychology, Montréal, Québec, Canada.
- Anderson, N. (2003). Applicant and recruiter reactions to new technology in selection: A critical review and agenda for future research. *International Journal of Selection and Assessment*, 11, 121–136. doi.org/10.1111/1468-2389.00235
- Anderson, N. R., Ahmed, S., & Costa, A. C. (2012). Applicant reactions in Saudi Arabia: Organizational attractiveness and core self-evaluation. *International Journal of Selection and Assessment*, 20, 197-208. https://doi.org/10.1111/j.1468-2389.2012.00592.x
- Armstrong, M. B., Ferrell, J. Z., Collmus, A. B., & Landers, R. N. (2016b). Correcting misconceptions about gamification of assessment: More than SJTs and badges. *Industrial and Organizational Psychology*, 9, 671-677. doi.org/10.1017/iop.2016.69

- Armstrong, M. B., Landers, R. N., & Collmus, A. B. (2016a). Gamifying recruitment, selection, training, and performance management: Game-thinking in human resource management. In D. Davis & H. Gangadharbatla (Eds.), *Handbook of research on trends in gamification* (pp.140-165). Hershey, PA: Information Science Reference. doi, 10.4018/978-1-4666-8651-9.ch007.
- Arthur, W., Doverspike, D., Muñoz, G. J., Taylor, J. E., & Carr, A. E. (2014). The use of mobile devices in high-stakes remotely delivered assessments and testing. *International Journal of Selection and Assessment*, 22, 113-123. doi.org/10.1111/ijsa.12062
- Barnett, E. (2012). Facebook passwords “fair game in job interviews.” *Telegraph*. Retrieved from: <http://www.telegraph.co.uk/technology/facebook/9162356/Facebook-passwords-fair-game-in-job-interviews.html>.
- Bauer, T., Truxillo, D., Tucker, J., Weathers, V., Bertolino, M., Erdogan, B., & Campion, M. (2006). Selection in the information age: The impact of privacy concerns and computer experience on applicant reactions. *Journal of Management*, 32, 601-621. doi.org/10.1177/0149206306289829
- Beaty, J. C., Nye, C. D., Borneman, M. J., Kantrowitz, T. M., Drasgow, F., & Grauer, E. (2011). Proctored versus unproctored Internet tests: Are unproctored noncognitive tests as predictive of job performance? *International Journal of Selection and Assessment*, 19, 1-10. doi.org/10.1111/j.1468-2389.2011.00529.x
- Becton, J. B., Walker, H. J., Gilstrap, J. B., & Schwager, P. H. (2019). Social media snooping on job applicants: The effects of unprofessional social media information on recruiter perceptions. *Personnel Review*, 48, 1261-1280. doi.org/10.1108/PR-09-2017-0278

- Berry, C. M., Cullen, M. J., & Meyer, J. M. (2014). Racial/ethnic subgroup differences in cognitive ability test range restriction: Implications for differential validity. *Journal of Applied Psychology, 99*, 21-37.
- Berry, C. M., & Zhao, P. (2015). Addressing criticisms of existing predictive bias research: Cognitive ability test scores still overpredict African Americans' job performance. *Journal of Applied Psychology, 100*, 162-179. .doi.org/10.1037/a0037615
- Black, S. L., Stone, D. L., & Johnson, A. F. (2015). Use of social networking websites on applicants' privacy. *Employee Responsibilities and Rights Journal, 27*, 115-159. doi 10.1007/s10672-014-9245-2
- Braddy, P. W., Meade, A. W., Michael, J. J., & Fleenor, J. W. (2009). Internet recruiting: Effects of website content features on viewers' perceptions of organizational culture. *International Journal of Selection and Assessment, 17*, 19-34. doi.org/10.1111/j.1468-2389.2009.00448.x
- Breaugh, J., & Starke, M. (2000). Research on employee recruitment: So many studies, so many remaining questions. *Journal of Management, 26*, 405–434. doi.org/10.1177/014920630002600303
- Brenner, F. S., Ortner, T. M., & Fay, D. (2016). Asynchronous video interviewing as a new technology in personnel selection: The applicant's point of view. *Frontiers in psychology, 7*, 863. doi.org/10.3389/fpsyg.2016.00863
- Brown, M. I., & Grossenbacher, M. A. (2017). Can you test me now? Equivalence of GMA tests on mobile and non-mobile devices. *International Journal of Selection and Assessment, 25*, 61-71. doi.org/10.1111/ijsa.12160
- Brown, V. & Vaughn, D. (2011). The writing on the (Facebook) wall: The use of social networking sites in hiring decisions. *Journal of Business Psychology, 26*, 219-225. doi10.1007/s10869-011-9221-x

Bruck-Lee, V., Lanz, J., Drew, E. N., Coughlin, C., Levine, P., Tuzinski, K., & Wrenn, K.

(2016). Examining Applicant Reactions to Different Media Types in Character-based Simulations for Employee Selection. *International Journal of Selection and Assessment*, 24, 77-91. doi.org/10.1111/ijsa.12132

Career Builder (2017). Number of Employers Using Social Media to Screen Candidates at All-Time High, Finds Latest CareerBuilder Study. Retrieved from:

<https://www.prnewswire.com/news-releases/number-of-employers-using-social-media-to-screen-candidates-at-all-time-high-finds-latest-careerbuilder-study-300474228.html>

Chamorro-Premuzic, T., Winsborough, D., Sherman, R. A., & Hogan, R. (2016). New talent signals: Shiny new objects or a brave new world?. *Industrial and Organizational Psychology*, 9, 621-640. doi:10.1017/iop.2016.6

Chapman, D. S., Uggerslev, K. L., & Webster, J. (2003). Applicant reactions to face-to-face and technology-mediated interviews: A field investigation. *Journal of applied psychology*, 88, 944-953. Doi: 10.1037/0021-9010.88.5.944

Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology*, 63, 83-117. doi.org/10.1111/j.1744-6570.2009.01163.x

Clark, L. A., & Roberts, S. J. (2010). Employer's use of social networking sites: A socially irresponsible practice. *Journal of Business Ethics*, 95, 507-525. doi 10.1007/s10551-010-0436-y.

Colella, A., Hebl, M., & King, E. (2017). One hundred years of discrimination research in the Journal of Applied Psychology: A sobering synopsis. *Journal of Applied Psychology*, 102, 500-513. doi.org/10.1037/apl0000084

- Coursaris, C. K., Hassanein, K., Head, M. M., & Bontis, N. (2012). The impact of distractions on the usability and intention to use mobile devices for wireless data services. *Computers in Human Behaviour*, 28, 1439-1449.
doi.org/10.1016/j.chb.2012.03.006
- Courtland, R. (2018). Bias detectives: the researchers striving to make algorithms fair. *Nature*, 558(7710), 357-357.
- Coyne, I., Warszta, T., Beadle, S., & Sheehan, N. (2005). The Impact of Mode of Administration on the Equivalence of a Test Battery: A Quasi-Experimental Design. *International Journal of Selection and Assessment*, 13, 220-224.
doi.org/10.1111/j.1468-2389.2005.00318.x
- De Goede, M. E., Van Vianen, A. E., & Klehe, U. C. (2011). Attracting applicants on the Web: PO fit, industry culture stereotypes, and website design. *International Journal of Selection and Assessment*, 19, 51-61. doi.org/10.1111/j.1468-2389.2010.00534.x
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2014). Du game design au gamefulness: définir la gamification. *Sciences du jeu*, (2). Retrieved from:
<https://journals.openedition.org/sdj/287>
- Dineen, B. R., Ling, J., Ash, S. R., & Del Vecchio, D. (2007). Aesthetic properties and message customization: navigating the dark side of web recruitment. *Journal of applied psychology*, 92, 356-372. doi: 10.1037/0021-9010.92.2.356
- Drouin, M., O'Connor, K. W., Schmidt, G. B., & Miller, D. A. (2015). Facebook fired: Legal perspectives and young adults' opinions on the use of social media in hiring and firing decisions. *Computers in Human Behaviour*, 46, 123-128.
doi.org/10.1016/j.chb.2015.01.011

EU FRA (2018). Big data, algorithms and discrimination. Retrieved from:

<https://fra.europa.eu/en/publication/2018/big-data-discrimination>

EU FRA (2017). Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights. Retrieved from:

https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-data-quality-and-ai_en.pdf

EU GDPR (2018). The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years. Retrieved from:

<https://eugdpr.org/>.

FAMA (2018a). Beyond Background Check: Ai-Powered Social Media Screening. Retrieved from: https://www.fama.io/social-media-background-check-b?gclid=Cj0KCQiA-c_iBRChARIsAGCOpB3_hQ42Nt8E4Osi9os0x5pWn7prY7n2PVWwe39TjHX6ZAVSTSBBFKAaApkYEALw_wcB.

FAMA (2018b). FAMA Legal Ebook. Retrieved from: <https://www.fama.io/fama-legal-ebook>.

Fernández, C., & Fernández, A. (2019). Ethical and Legal Implications of AI Recruiting Software. *ERCIM NEWS*, 116, 22-23. Retrieved from: <https://ercim-news.ercim.eu/images/stories/EN116/EN116-web.pdf#page=22>

García, D. A., García, J. C. A., Izquierdo, A. L. G., & Rodríguez, J. (2019). LinkedIn “Big Four”: Job Performance Validation in the ICT Sector. *Revista de psicología del trabajo y de las organizaciones= Journal of work and organizational psychology*, 35, 53-64.

Georgiou, K., Gouras, A., & Nikolaou, I. (2019). Gamification in employee selection: The development of a gamified assessment. *International Journal of Selection and Assessment*, 27, 91-103. doi.org/10.1111/ijsa.12240

- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18, 694-734.
doi.org/10.5465/amr.1993.9402210155
- Gorman, C. A., Robinson, J., & Gamble, J. S. (2018). An investigation into the validity of asynchronous web-based video employment-interview ratings. *Consulting Psychology Journal: Practice and Research*, 70, 129-146. doi.org/10.1037/cpb0000102
- Griffen, A. (2018). Independent: Facebook is in crisis over data abuse scandal and Harvesting of information was rife says academic at heart of it. Retrieved from:
<https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-data-abuse-scandal-latest-cambridge-analytica-kogan-a8320066.html>
- Gruca, R. A., & Goldberg, L. R. (2007). The comparative validity of 11 modern personality inventories: Predictions of behavioural acts, informant reports, and clinical indicators. *Journal of Personality Assessment*, 89, 167-187.
doi.org/10.1080/00223890701468568
- Guardian UK (2017) *How white engineers built racist code – and why it's dangerous for black people*. Retrieved from:
<https://www.theguardian.com/technology/2017/dec/04/racist-facial-recognition-white-coders-black-people-police>
- Guchait, P., Ruetzler, T., Taylor, J., & Toldi, N. (2014). Video interviewing: A potential selection tool for hospitality managers—A study to understand applicant perspective. *International Journal of Hospitality Management*, 36, 90-100.
doi.org/10.1016/j.ijhm.2013.08.004
- Guo, J., & Drasgow, F. (2010). Identifying cheating on unproctored internet tests: The Z-test and the likelihood ratio test. *International Journal of Selection and Assessment*, 18, 351-364. doi.org/10.1111/j.1468-2389.2010.00518.x

- Hiemstra, A. M., Oostrom, J. K., Deros, E., Serlie, A. W., & Born, M. P. (2019). Applicant perceptions of initial job candidate screening with asynchronous job interviews: Does personality matter? *Journal of Personnel Psychology, 18*, 138-147.
doi.org/10.1027/1866-5888/a000230
- Jackson, D. J., LoPilato, A. C., Hughes, D., Guenole, N., & Shalfrooshan, A. (2017). The internal structure of situational judgement tests reflects candidate main effects: Not dimensions or situations. *Journal of Occupational and Organizational Psychology, 90*, 1-27. doi.org/10.1111/joop.12151
- Jeske, D., & Shultz, K. S. (2019). Social media screening and content effects: implications for job applicant reactions. *International Journal of Manpower, 40*, 73-86.
doi.org/10.1108/IJM-06-2017-0138
- Karim, M. N., Kaminsky, S. E., & Behrend, T. S. (2014). Cheating, reactions, and performance in remotely proctored testing: An exploratory experimental study. *Journal of Business and Psychology, 29*, 555-572. doi: 10.1007/s10869-014-9343-z
- King, D. D., Ryan, A. M., Kantrowitz, T., Grelle, D., & Dainis, A. (2015). Mobile Internet testing: An analysis of equivalence, individual differences, and reactions. *International Journal of Selection and Assessment, 23*, 382-394.
doi.org/10.1111/ijsa.12122
- Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the Age of Algorithms. *Journal of Legal Analysis, 10*. doi.org/10.1093/jla/laz001
- Kluemper, D. H., & Rosen, P. A. (2009). Future employment selection methods: Evaluating social networking Web sites. *Journal of Managerial Psychology, 24*, 567-580.
doi.org/10.1108/02683940910974134
- Kluemper, D. H., Rosen, P. A., & Mossholder, K. W. (2012). Social networking Websites, personality ratings, and the organizational context: More than meets the eye? *Journal*

of Applied Social Psychology, 42, 1143-1172. doi.org/10.1111/j.1559-1816.2011.00881.x

Konradt, U., Warszta, T., & Ellwart, T. (2013). Fairness perceptions in Web-based selection: Impact on applicants' pursuit intentions, recommendation intentions, and intentions to reapply. *International Journal of Selection and Assessment*, 21, 155-169. doi.org/10.1111/ijsa.12026

Kraichy, D., & Chapman, D. S. (2014). Tailoring web-based recruiting messages: Individual differences in the persuasiveness of affective and cognitive messages. *Journal of Business and Psychology*, 29, 253-268. doi: 10.1007/s10869-013-9311-z

Lambrecht, A., & Tucker, C. (2019). Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Management Science*. 65, 2947-3448. doi.org/10.1287/mnsc.2018.3093

Landers, R. N., & Sackett, P. R. (2012). Offsetting performance losses due to cheating in unproctored internet-based testing by increasing the applicant pool. *International Journal of Selection and Assessment*, 20, 220-228. doi.org/10.1111/j.1468-2389.2012.00594.x

Langer, M., König, C. J., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment*, 25, 371-382. doi.org/10.1111/ijsa.12191

Laumer, S., Eckhardt, A., & Weitzel, T. (2012). Online gaming to find a new job—examining job seekers' intention to use serious games as a self-assessment tool. *German Journal of Human Resource Management*, 26, 218-240. doi.org/10.1177/239700221202600302

- Le Corff, Y., Gingras, V., & Busque-Carrier, M. (2017). Equivalence of unproctored internet testing and proctored paper-and-pencil testing of the Big Five. *International Journal of Selection and Assessment*, 25, 154-160. doi.org/10.1111/ijsa.12168
- Leggatt, H. (2010). Germany bans Facebook as employee screening tool. *BizReport*.
<http://www.bizreport.com/2010/09/germany-bans-facebook-as-employee-screening-tool.html>.
- Levashina, J., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology*, 67, 241-293. doi.org/10.1111/peps.12052
- Lievens, F., & Burke, E. (2011). Dealing with the threats inherent in unproctored Internet testing of cognitive ability: Results from a large-scale operational test program. *Journal of Occupational and Organizational Psychology*, 84, 817-824.
doi.org/10.1348/096317910X522672
- Lievens, F., De Corte, W., & Westerveld, L. (2015). Understanding the building blocks of selection procedures: Effects of response fidelity on performance and validity. *Journal of Management*, 41, 1604-1627. doi.org/10.1177/0149206312463941
- Lievens, F., & Harris, M.M. (2003). Research on Internet Recruiting and Testing: Current Status and Future Directions. In C.L. Cooper & I.T. Robertson (Eds.) *International Review of Industrial and Organizational Psychology*, vol. 16. (pp. 131-165).
Chicester: John Wiley & Sons, Ltd.
- Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of applied psychology*, 91, 1181-1188. doi.org/10.1037/0021-9010.91.5.1181

- Madera, J. M. (2012). Using social networking websites as a selection tool: The role of selection process fairness and job pursuit intentions. *International Journal of Hospitality Management*, 31, 1276–1282. doi:10.1016/j.ijhm.2012.03.008.
- McCarthy, J. M., Bauer, T. N., Truxillo, D. M., Anderson, N. R., Costa, A. C., & Ahmed, S. M. (2017). Applicant Perspectives During Selection: A Review Addressing “So What?,” “What’s New?,” and “Where to Next?”. *Journal of Management*, 43, 1693–1725. <https://doi.org/10.1177/0149206316681846>
- Meade, A. W., Michels, L. C., & Lautenschlager, G. J. (2007). Are Internet and paper-and-pencil personality tests truly comparable? An experimental design measurement invariance study. *Organizational Research Methods*, 10, 322-345. doi.org/10.1177/1094428106289393.
- Morelli, N. A., Mahan, R. P., & Illingworth, A. J. (2014). Establishing the measurement equivalence of online selection assessments delivered on mobile versus nonmobile devices. *International Journal of Selection and Assessment*, 22, 124-138. doi.org/10.1111/ijsa.12063
- Morelli, N., Potosky, D., Arthur, W., & Tippins, N. (2017). A call for conceptual models of technology in IO psychology: An example from technology-based talent assessment. *Industrial and Organizational Psychology*, 10, 634-653. doi.org/10.1017/iop.2017.70
- Moscato, S., Salgado, J. F., & Anderson, N. R. (2017). How Do I Get a Job, What Are They Looking For? In N. Chmiel, F. Fraccaroli, & M. Sverke (Eds.), *An Introduction to Work and Organizational Psychology: An International Perspective*. London: Wiley/Blackwell.

- Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., & Velasquez, R. (2004). Psychological testing on the Internet: new problems, old issues. *American Psychologist*, 59, 150 -162. doi.org/10.1037/0003-066X.59.3.150.
- Nikolaou, I. (2014). Social networking web sites in job search and employee recruitment. *International Journal of Selection and Assessment*, 22, 179-189. doi.org/10.1111/ijsa.12067.
- Nikolaou, I., Bauer, T. N., & Truxillo, D. M. (2015). Applicant reactions to selection methods: An overview of recent research and suggestions for the future. In I. Nikolaou, & J. Oostrom (Eds). *Employee recruitment, selection, and assessment. Contemporary Issues for Theory and Practice* (pp 80-96). Hove, East Sussex: Routledge.
- Nikolaou, I., Georgiou, K. & Kotsasarlidou, V. (2019) The Exploring the relationship of a game-based assessment with performance. *Spanish Journal of Psychology*....
- Oh, I.-S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, 96, 762-773. doi.org/10.1037/a0021832
- Ötting, S. K., & Maier, G. W. (2018). The importance of procedural justice in human-machine interactions: Intelligent systems as new decision agents in organizations. *Computers in Human Behaviour*, 89, 27-39. <https://doi.org/10.1016/j.chb.2018.07.022>.
- Patterson, F., Ashworth, V., Zibarras, L., Coan, P., Kerrin, M., & O'Neill, P. (2012). Evaluations of situational judgement tests to assess non-academic attributes in selection. *Medical education*, 46, 850-868. doi.org/10.1111/j.1365-2923.2012.04336.x

- Pfieffelman, B., Wagner, S. H., & Libkuman, T. (2010). Recruiting on corporate web sites: Perceptions of fit and attraction. *International Journal of Selection and Assessment*, 18, 40-47. doi.org/10.1111/j.1468-2389.2010.00487.x
- Ployhart, R. E. (2006). Staffing in the 21st century: New challenges and strategic opportunities. *Journal of management*, 32, 868-897. doi.org/10.1177/0149206306293625
- Ployhart, R. E., Schmitt, N., & Tippins, N. T. (2017). Solving the Supreme Problem: 100 years of selection and recruitment at the Journal of Applied Psychology. *Journal of Applied Psychology*, 102, 291-304. doi.org/10.1037/apl0000081
- Ployhart, R. E., Weekley, J. A., Holtz, B. C., & Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: are personality, biodata, and situational judgment tests comparable? *Personnel Psychology*, 56, 733-752. doi.org/10.1111/j.1744-6570.2003.tb00757.x
- Potosky, D. (2008). A conceptual framework for the role of the administration medium in the personnel assessment process. *Academy of Management Review*, 33, 629-648. doi.org/10.5465/amr.2008.32465704
- Potosky, D., & Bobko, P. (2004). Selection testing via the Internet: Practical considerations and exploratory empirical findings. *Personnel Psychology*, 57, 1003-1034. doi.org/10.1111/j.1744-6570.2004.00013.x
- Roth, P. L., Bobko, P., Van Iddekinge, C. H., & Thatcher, J. B. (2016). Social media in employee-selection-related decisions a research agenda for uncharted territory. *Journal of Management*. 42, 1-30. doi: 10.1177/0149206313503018.
- Roth, P. L., Le, H., Oh, I. S., Van Iddekinge, C. H., Buster, M. A., Robbins, S. B., & Campion, M. A. (2014). Differential validity for cognitive ability tests in employment

- and educational settings: Not much more than range restriction?. *Journal of Applied Psychology*, 99, 1-20. doi.org/10.1037/a0034377
- Roth, P. L., Le, H., Oh, I. S., Van Iddekinge, C. H., & Robbins, S. B. (2017). Who ru?: On the (in) accuracy of incumbent-based estimates of range restriction in criterion-related and differential validity research. *Journal of Applied Psychology*, 102, 802-828. <http://dx.doi.org/10.1037/apl0000193>.
- Roulin, N. (2014). The influence of employers' use of social networking websites in selection, online self-promotion, and personality on the likelihood off faux pas postings. *International Journal of Selection and Assessment*, 22, 80-8. doi.org/10.1111/ijsa.12058
- Roulin, N., & Levashina, J. (2016). Impression management and social media profiles. In R. N. Lander & G. B. Schmidt (Eds). *Social Media in Employee Selection and Recruitment* (pp. 223-248). AG Switzerland: Springer International Publishing.
- Ryan, A. M., & Derous, E. (2019). The unrealized potential of technology in selection assessment. *Journal of Work and Organizational Psychology*. Advance online publication.
- Sajjadiani, S., Sojourner, A. J., Kammeyer-Mueller, J. D., & Mykerezzi, E. (2019). Using machine learning to translate applicant work history into predictors of performance and turnover. *Journal of Applied Psychology*. Advance online publication.
- Salgado, J. F., & Moscoso, S. (2003). Internet-based personality testing: Equivalence of measures and assesses' perceptions and reactions. *International Journal of selection and assessment*, 11, 194-205. doi.org/10.1111/1468-2389.00243
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological bulletin*, 124, 262-274.

- Schmidt, F. L., Oh, I. S., & Shaffer, J. A. (2016). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 100 Years... Fox School of Business Research Paper. Retrieved from <https://www.testingtalent.net/wp-content/uploads/2017/04/2016-100-Yrs-Working-Paper-on-Selection-Methods-Schmit-Mar-17.pdf>
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8, e73791. doi.org/10.1371/journal.pone.0073791. Retrieved from <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0073791&type=printable>
- Scott, J. C., Bartram, D., & Reynolds, D. H. (2018). Next Generation Technology-Enhanced Assessment. Cambridge, UK: Cambridge University Press.
- Sears, G. J., Zhang, H. H., Wiesner, W. D., Hackett, R. D., & Yuan, Y. (2013). A comparative assessment of videoconference and face-to-face employment interviews. *Management Decision*, 51, 1733-1752. doi.org/10.1108/MD-09-2012-0642.
- Selden, S., & Orenstein, J. (2011). Government e-recruiting web sites: The influence of e-recruitment content and usability on recruiting and hiring outcomes in US state governments. *International Journal of Selection and Assessment*, 19, 31-40. doi.org/10.1111/j.1468-2389.2011.00532.x
- Silvester, J., & Anderson, N. (2003). Technology and discourse: A comparison of face-to-face and telephone employment interviews. *International Journal of Selection and Assessment*, 11, 206-214. doi.org/10.1111/1468-2389.00244

Sinar, E. F., Reynolds, D. H., & Paquet, S. L. (2003). Nothing but 'Net? Corporate image and web-based testing. *International Journal of Selection and Assessment*, 11, 150-157.

doi.org/10.1111/1468-2389.00238

Social Intelligence (2018). Who we are. Retrieved from: <https://info.socialintel.com/about-us>

Steiner, K. (2017). Bad Candidate Experience Cost Virgin Media \$5M Annually – Here is How They Turned That Around. Retrieved from: <https://business.linkedin.com/talent-solutions/blog/candidate-experience/2017/bad-candidate-experience-cost-virgin-media-5m-annually-and-how-they-turned-that-around>

Stinson, J. (2014). Password Protected: States Pass Anti-Snooping Laws. Retrieved from: <http://www.govtech.com/gov-experience/Password-Protected-States-Pass-Anti-Snooping-Laws.html> .

Stoughton, J. W., Thompson, L. F., & Meade, A. W. (2015). Examining applicant reactions to the use of social networking websites in pre-employment screening. *Journal of Business and Psychology*, 30, 73-88. doi: 10.1007/s10869-013-9333-6

Straus, S. G., Miles, J. A., & Levesque, L. L. (2001). The effects of videoconference, telephone, and face-to-face media on interviewer and applicant judgments in employment interviews. *Journal of management*, 27, 363-381. doi: doi.org/10.1177/014920630102700308

Suen, H. Y., Chen, M. Y. C., & Lu, S. H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes?. *Computers in Human Behaviour*, 98, 93-101. doi.org/10.1016/j.chb.2019.04.012

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, 29, 24-54. DOI: 10.1177/0261927X09351676.

- Templer, K. J., & Lange, S. R. (2008). Internet testing: Equivalence between proctored lab and unproctored field conditions. *Computers in Human Behaviour*, 24, 1216-1228. doi.org/10.1016/j.chb.2007.04.006
- Tendeiro, J. N., Meijer, R. R., Schakel, L., & Maij-de Meij, A. M. (2013). Using cumulative sum statistics to detect inconsistencies in unproctored Internet testing. *Educational and Psychological Measurement*, 73, 143-161. doi: 10.1177/0013164412444787
- Tippins, N. T. (2015). Technology and assessment in selection. *Annual Review of Organizational Psychology and Organizational Behaviour*, 2, 551-582. doi.org/10.1146/annurev-orgpsych-031413-091317.
- Tippins, N. T., Beaty, J., Drasgow, F., Gibson, W. M., Pearlman, K., Segall, D. O., & Shepherd, W. (2006). Unproctored internet testing in employment settings. *Personnel Psychology*, 59, 189-225. doi.org/10.1111/j.1744-6570.2006.00909.x
- Truxillo, D., Bauer, T., McCarthy, J., Anderson, N., & Ahmed, S. (2018). *Applicant perspectives on employee selection systems*. In N. R. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran, C. (Eds.). *Handbook of Industrial, Work & Organizational Psychology*. London: Sage.
- van Esch, P., Black, J. S., & Ferolie, J. (2019). Marketing AI recruitment: the next phase in job application and selection. *Computers in Human Behaviour*, 90, 215-222. doi.org/10.1016/j.chb.2018.09.009
- Van Hoya, G. (2013). Word-of-Mouth as a Recruitment Source: An Integrative Model. In D. M. Cable & K. Y. T. Yu (Eds.), *Oxford Handbook of Recruitment*. New York: Oxford Press. doi: 10.1093/oxfordhb/9780199756094.013.023
- Van Iddekinge, C. H., Lanivich, S. E., Roth, P. E., & Junco, E. (2016). Social media for selection? Validity and adverse impact potential of a Facebook-based assessment. *Journal of Management* 42, 1811-1835. doi:10.1177/0149206313515524.

- Vicknair, J., Elkersh, D., Yancey, K., & Budden, M. C. (2010). The use of social networking websites as a recruiting tool for employers. *American Journal of Business Education, 3*, 7-12.
- Weekley, J. A., Hawkes, B., Guenole, N., & Ployhart, R. E. (2015). Low-fidelity simulations. *Annual Review of Organizational Psychology and Organizational Behaviour, 2*, 295-322. doi.org/10.1146/annurev-orgpsych-032414-111304.
- Woods, S. A. (2018) Assessment in the digital age: some challenges for test developers and users. *Assessment and Development Matters*. Retrieved from:
<https://www.astonassessments.com/wp-content/uploads/2018/05/Assessment-in-the-Digital-Age-formatted-for-downloads.pdf>
- Woods, S. A., & Anderson, N. R. (2016). Toward a periodic table of personality: Mapping personality scales between the five-factor model and the circumplex model. *Journal of Applied Psychology, 101*, 582 -604. <http://dx.doi.org/10.1037/apl0000062>
- Woods, S. A. & West, M. A. (2019). *The Psychology of Work and Organizations*. CENGAGE: London.
- Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (2018). Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard? *Philosophy & Technology, 1*-23. doi.org/10.1007/s13347-018-0330-6
- Zibarras, L. D., Patterson, F., Holmes, J., Flaxman, C., & Kubacki, A. (2018). An exploration of applicant perceptions of asynchronous video MMIs in medical selection. *MedEdPublish, 7*, 64. 10.15694/mep.2018.0000285.1

Table 1. Ten practice-oriented questions for future research on Digital Selection Procedures to address.

1. What are appropriate and effective benchmarks for evaluating the construct validity of DSPs?
2. How can validity evidence be applied most effectively in the rapid design, improvement and deployment of DSPs?
3. What are the criterion validities of different forms of DSPs?
4. Do DSPs outperform non-digital selection procedures in terms of validity in any specific situations?
5. What are the incremental criterion validities attained through using a combination of DSPs and non-digital methods?
6. What are the effects of potential new forms of adverse impact and bias in the use of DSPs, and in what ways do they exacerbate or ameliorate unfairness concerns in selection?
7. How do DSPs shape applicant reactions toward the hiring organization?
8. How should the basis for digital selection decisions be presented to applicants to ensure transparency?
9. What are the ethical and privacy considerations that should be considered by the organizations to enhance applicant reactions to DSPs?
10. Are there generational or other relevant demographic differences in applicant reactions to DSPs?